Wages, Productivity and Aging*

Benoit Dostie†

January 20th 2006

Abstract

In this article, we estimate age based wages and productivity differentials using linked employer-employee Canadian data from the Workplace and Employee Survey 1999-2002. Data on the firm side is used to estimate production functions taking into account the age profile of the firm's workforce. Data on the workers' side is used to estimate wage equations that also depend on age. Results show concave age-wage and age-productivity profiles and wage-productivity comparisons show that the productivity of workers aged 55 and more decreases faster than their wages.

1 Introduction

Wage differentials based on different levels of schooling or experience are well documented in the labor economics literature. These are generally interpreted as productivity differentials based on an investment model of human capital. These models generally predict that wages increase in the early stages of a

*PRELIMINARY AND INCOMPLETE VERSION. I would like to thank Rob Clark, Rajshri Jayaraman, André Léonard and Étienne Wasmer for their comments. I also thank Marie-Pierre Pelletier for her excellent research assistance. I acknowledge funding from Human Resources and Skills Development Canada. The usual caveats apply.

†Institute of applied economics, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7; email: benoit.dostie@hec.ca; fax: 514-340-6469; phone : 514-340-6453; CIRANO, CIRPÉE and IZA.

career until they reach a plateau, afterwhich, they start do diminish due to human capital depreciation (Mincer (1974)).

There are many reasons, however, to doubt that wage differentials always correspond to differences in productivity. Among alternative explanations, one could include incentive-compatible wages (Lazear (1979)), forced saving mechanisms (Frank and Hutchens (1993) and Loewenstein and Sicherman (1991)), efficiency wages (Akerlof and Yellen (1986)), minimization of turnover related costs (Salop and Salop (1976)), specific training (Hutchens (1989)) or wage discrimination.[1]

In one of the first detailed empirical studies on this topic, Medoff and Abraham (1980) find evidence that the wages of older workers might be higher than their productivity. Corroborating evidence has been found in numerous countries and for different professions using a variety of methodologies, including Oster and Hamermesh (1998) for economists, Kotlikoff and Gokhale (1992) for workers in the U.S. manufacturing industries, Fair (1994) for American athletes, Crépon, Deniau, and Pérez-Duarte (2003) for France[2] and Haegeland and Klette (1999) for Norway. However, many other studies find the opposite result, for example Mitchell (1990), Smith (1991), Hellerstein and Neumark (1995) and Hellerstein, Neumark, and Troske (1999). Skirbekk (2003) summarizes the evidence on the topic and nonetheless concludes that individual productivity decreases rapidly after 50

In this article, we estimate age-specific wage and productivity differentials using Canadian data from the Workplace and Employee Survey (WES) 1999-2002. The survey is designed to be representative of all firms operating in Canada and contains detailed information on the each firm's production pro-

[1]Discrepancies between wages and productivity could also arise due to labor market imperfections (for example due to minimum wages laws and trade unions, or oligopsonistic wage-setting (Acemogly and Pischke (1999)), etc.).

[2]Aubert and Crépon (2004) use a slightly different model and get completely different results.

cess, organizational practices (and changes in such practices), and human resources policies. Since the survey is linked, there is no need to assign workers to firms using statistical matching methods like Hellerstein, Neumark, and Troske (1999). Also, because the survey is linked, we are able to obtain an external value for a worker's productivity, independent of his or her wage.

More specifically, we use data at the workplace level to estimate production functions taking into account the age composition of the firm's workforce and use data at the employee level to estimate wage equations distinguishing workers based on their age.

Our methodology is thus similar to Hellerstein, Neumark, and Troske (1999) and Aubert and Crépon (2004). However, we improve on their methodologies for estimating wage equations by taking into account both individual and firm unobserved heterogeneity using a mixed model of wage determination (as suggested by Abowd and Kramarz (1999b)). We also control for unobserved time-varying productivity shocks in the production function using a method suggested by Levinsohn and Petrin (2003).

We find that wage profiles are not sensitive to the inclusion of unobserved heterogeneity at the workplace and worker levels. However, wage profiles are different when estimating wage equations at the worker level compared to wage profiles obtained by estimating a model of the determinants of the firm's payroll. We also find productivity profiles to be steeper once unobserved productivity shocks are controlled for. Finally, while we find concave profiles for both wage and productivity, our results also show that productivity is diminishing faster than wages for workers aged 55 and over.

The plan of rest of the paper is as follows. We first describe our methodology in section 2 and present the data and some descriptive statistics in the following section. We describe the results in section 4 and conclude briefly in section 5.

All tables are in the appendix.

2 Methodology

Our methodology improves on previous work in two ways (1) we take into account firm unobserved heterogeneity (in addition to worker unobserved heterogeneity) in the estimation of the wage equation and (2) we also take into account unobserved time-varying productivity shocks using an estimation method suggested by Levinsohn and Petrin (2003) in the estimation of the production function. We describe both models in the following subsections.

2.1 Production function

In order to estimate age-productivity profiles, first consider a Cobb-Douglas production function

$$\log Q_{jt} = \alpha \log L_{jt}^{A} + \beta \log K_{jt} + u_{jt} \tag{1}$$

where Q is the value added by firm j at time t, L^A is an aggregate function of different types of workers[3], K is the capital stock and u the error term. For each firm, we observe a representative sample of workers. In what follows, we use this sample to distinguish different types of workers based on age, education and gender[4].

Let L_{jtk} be the number of workers of type k in firm j at time t, and ϕ_k be their productivity[5]. If we assume that workers of each type are perfectly

[3] It would be better to use a measure of hours worked but we do not have this information at the firm level in the data set.

[4] It would be interesting to disaggregate the sample based on some other dimensions (like occupation for example). However, given the relatively small number of workers that was sampled form each firm, we stick with these 3 characteristics in the analysis that follows.

[5] We assume that a worker has the same marginal product accross firm.

substitutable, we can write

$$L^A_{jt} = \sum_0^K \lambda_k L_{jtk} = \lambda_{j0} L_{jt} + \sum_1^K (\lambda_{jk} - \lambda_{j0}) L_{itk} \tag{2}$$

where L_{jt} is the total number of workers in the firm λ_0 the productivity of the reference category of workers. We can rewrite equation (2) as

$$\log L^A_{jt} = \log \lambda_0 + \log L_{jt} + \log \left(1 + \sum_1^K \left(\frac{\lambda_k}{\lambda_0} - 1\right) P_{jkt}\right) \tag{3}$$

where P_{jkt} is the ratio of the number of workers of type k over the total number of employees. We then write the production function as

$$\begin{aligned} \log Q_{jt} &= \alpha \log \lambda_0 + \alpha \log L_{jt} + \\ &+\alpha \log \left(1 + \sum_1^K \left(\frac{\lambda_k}{\lambda_0} - 1\right) P_{jkt}\right) + \beta \log K_{jt} + u_{jt} \end{aligned} \tag{4}$$

As Hellerstein, Neumark, and Troske (1999), we distinguish three age groups: less then 35, between 35 and 55, and over 55. As to education, we distinguish workers based on whether they have a graduate diploma or not. Therefore, workers are thus separated in 8 categories (men and women (H and F); young, middle age or old (J, M and V); with or without a diploma (S, A)). If we take young male workers without a graduate diploma as our reference category, we can write:

$$\log L_{jt}^{A} = \log\lambda_0 + \log L_{jt} + \\ + \log\left(\begin{array}{c} 1 + \gamma_{HJA}\frac{L_{HJAjt}}{L_{jt}} + \gamma_{HMS}\frac{L_{HMSjt}}{L_{jt}} + \gamma_{HMA}\frac{L_{HMAjt}}{L_{jt}} \\ \gamma_{HVS}\frac{L_{HVSjt}}{L_{jt}} + \gamma_{HVA}\frac{L_{HVAjt}}{L_{jt}} + \\ \gamma_{FJS}\frac{L_{FJSjt}}{L_{jt}} + \gamma_{FJA}\frac{L_{FJAjt}}{L_{jt}} + \\ \gamma_{FMS}\frac{L_{FMSjt}}{L_{jt}} + \gamma_{FMA}\frac{L_{FMAjt}}{L_{jt}} + \\ \gamma_{FVS}\frac{L_{FVSjt}}{L_{jt}} + \gamma_{FVA}\frac{L_{FVAjt}}{L_{jt}} + \end{array}\right) \quad (5)$$

where γ equal $(\lambda/\lambda_0 - 1)$. Since $\log(1+x) \simeq x$, we can approximate this by

$$\begin{aligned} \log L_{jt}^{A} = {} & \log\lambda_0 + \log L_{jt} + \gamma_{HJA}\frac{L_{HJAjt}}{L_{jt}} + \\ & \gamma_{HMS}\frac{L_{HMSjt}}{L_{jt}} + \gamma_{HMA}\frac{L_{HMAjt}}{L_{jt}} + \\ & \gamma_{HVS}\frac{L_{HVSjt}}{L_{jt}} + \gamma_{HVA}\frac{L_{HVAjt}}{L_{jt}} + \\ & \gamma_{FJS}\frac{L_{FJSjt}}{L_{jt}} + \gamma_{FJA}\frac{L_{FJAjt}}{L_{jt}} + \\ & \gamma_{FMS}\frac{L_{FMSjt}}{L_{jt}} + \gamma_{FMA}\frac{L_{FMAjt}}{L_{jt}} + \\ & \gamma_{FVS}\frac{L_{FVSjt}}{L_{jt}} + \gamma_{FVA}\frac{L_{FVAjt}}{L_{jt}} \end{aligned} \quad (6)$$

We call this specification the "complete" model. If we impose the following restrictions: $\gamma_{HJA} = \gamma_A$, $\gamma_{HMS} = \gamma_M$, $\gamma_{HMA} = \gamma_M \cdot \gamma_A$, $\gamma_{HVS} = \gamma_V$, $\gamma_{HVA} = \gamma_V \cdot \gamma_A$, $\gamma_{FJS} = \gamma_F$, $\gamma_{FJA} = \gamma_F \cdot \gamma_A$, $\gamma_{FMS} = \gamma_F \cdot \gamma_M$, $\gamma_{FMA} = \gamma_F \cdot \gamma_M \cdot \gamma_A$, $\gamma_{FVS} = \gamma_F \cdot \gamma_V$, $\gamma_{FVA} = \gamma_F \cdot \gamma_V \cdot \gamma_A$, we can write a more parsimonious specification as

$$\log L_{jt}^{A} = \log\lambda_0 + \log L_{jt} + \gamma_F\frac{L_{Fjt}}{L_{jt}} + \gamma_M\frac{L_{Mjt}}{L_{jt}} + \gamma_V\frac{L_{Vjt}}{L_{jt}} + \gamma_A\frac{L_{Ajt}}{L_{jt}} \quad (7)$$

We call this last specification the "restricted" model. Substituting (2) in (4) gives us the restricted model:

$$\begin{aligned}\log Q_{jt} \quad \simeq \quad & \beta_0 + \alpha \log L_{jt} + \beta \log K_{jt} + \alpha\gamma_F \frac{L_{Fjt}}{L_{jt}} + \alpha\gamma_M \frac{L_{Mjt}}{L_{jt}} + \\ & \alpha\gamma_V \frac{L_{Vjt}}{L_{jt}} + \alpha\gamma_A \frac{L_{Ajt}}{L_{jt}} + +\delta' Z_{jt} + \ u_{jt} \end{aligned} \tag{8}$$

and the complete model follows from the substitution of (6) in (4) :

$$\begin{aligned}\log Q_{jt} \quad \simeq \quad & \beta_0 + \alpha \log L_{jt} + \beta \log K_{jt} + \alpha\gamma_{HJA} \frac{L_{HJAjt}}{L_{jt}} + \\ & \alpha\gamma_{HMS} \frac{L_{HMSjt}}{L_{jt}} + \alpha\gamma_{HMA} \frac{L_{HMAjt}}{L_{jt}} + \\ & \alpha\gamma_{HVS} \frac{L_{HVSjt}}{L_{jt}} + \alpha\gamma_{HVA} \frac{L_{HVAjt}}{L_{jt}} + \\ & \alpha\gamma_{FJS} \frac{L_{FJSjt}}{L_{jt}} + \alpha\gamma_{FJA} \frac{L_{FJAjt}}{L_{jt}} + \\ & \alpha\gamma_{FMS} \frac{L_{FMSjt}}{L_{jt}} + \alpha\gamma_{FMA} \frac{L_{FMAjt}}{L_{jt}} + \\ & \alpha\gamma_{FVS} \frac{L_{FVSjt}}{L_{jt}} + \alpha\gamma_{FVA} \frac{L_{FVAjt}}{L_{jt}} + \delta' Z_{jt} + u_{jt} \end{aligned} \tag{9}$$

where β_0 is a constant term that incorporates $\alpha \log \lambda_0$, Z_{jt} is a matrix of observed firm characteristics including the organizational practices of the firm and δ is a vector of parameters.

Note that coefficient estimates of equations (8) and (9) will be biased if input choices in the production function are correlated to unobserved productivity shocks. Profit maximizing firm will respond to a positive shock by increasing production which requires more input. In a similar manner, negative productivity shocks will lead firms to lower their production level. To correct for endogenous input choices, we use a two-stages estimation method suggested by Levinsohn and Petrin (2003). Their method requires the use of intermediate inputs as proxies, arguing that intermediates may respond more smoothly to productivity shocks. They work with a modification of equation (**??**) where[6]

$$\begin{aligned}\log Q_{jt} &= \beta_0 + \alpha \log L^A_{jt} + \beta \log K_{jt} + u_{jt} \\ \log Q_{jt} &= \beta_0 + \alpha \log L^A_{jt} + \beta \log K_{jt} + \omega_{jt} + \eta_{jt} \\ &= \alpha \log L^A_{jt} + \phi(K_{jt}, M_{jt}) + \eta_{jt} \qquad (10)\end{aligned}$$

where the error term u_{jt} has been expressed as a productivity shock ω_{jt} to which input choices might be correlated and an orthogonal residual η_{jt}. M_{jt} represents an intermediate input. Substituting a third order polynomial approximation in K_{jt} and M_{jt} in place of $\phi(K_{jt}, M_{jt})$ makes it possible to consistently estimate parameters of the value-added function using OLS as

$$\log Q_{jt} = \delta_0 + \alpha \log L^A_{jt} + \sum_{o=0}^{3} \sum_{p=0}^{3} \delta_{ij} K^o_{jt} M^p_{jt} + \eta_{jt}$$

where β_0 is not separately identified from the intercept $\phi_t (K_{jt}, M_{jt})$. This completes the first stage of the estimation routine, from which an estimate of β_l, and an estimate of ϕ (up to the intercept) are available.

The second stage of the routine identifies the coefficient β_k. It begins by computing the estimated value for ϕ_t using

$$\begin{aligned}\hat{\phi}_{jt} &= \log \hat{Q}_{jt} - \hat{\alpha} \log L^A_{jt} \\ &= \hat{\delta}_0 + \sum_{o=0}^{3} \sum_{p=0}^{3} \hat{\delta}_{ij} K^o_{jt} M^p_{jt} - \hat{\alpha} \log L^A_{jt}\end{aligned}$$

For any candidate value β^*, we can compute (up to a scalar constant) a prediction for ω_{jt} for all periods t using

$$\hat{\omega}_{jt} = \hat{\phi}_{jt} - \beta^* K_{jt}$$

[6]The exposition that follows draws heavily from Levinsohn and Petrin (2003).

Using these values, a consistent (nonparametric) approximation to $E[\omega_{jt}|\omega_{jt-1}]$ is given by the predicted values from the regression

$$\hat{\omega}_{jt} = \gamma_0 + \gamma_1\omega_{jt-1} + \gamma_2\omega_{jt-1}^2 + \gamma_3\omega_{jt-1}^3 + \epsilon_t$$

which we call $E[\omega_{jt}|\omega_{jt-1}]$

Given α, β^*, and $E[\omega_{jt}|\omega_{jt-1}]$, Levinsohn and Petrin (2003) write the sample residual of the production function as

$$\eta_{jt} = \log Q_{jt} - \hat{\alpha}\log L_{jt}^{A} - \beta^* K_{jt} - E[\omega_{jt}|\omega_{jt-1}]$$

and the estimates for β is obtained as the solution to

$$\min_{\beta}\sum_{t}\left(\log Q_{jt} - \hat{\alpha}\log L_{jt}^{A} - \beta^* K_{jt} - E[\omega_{jt}|\omega_{jt-1}]\right)^2$$

2.2 Wage equations

Turning to the estimation of the relationship between age and wages, it is possible to use two approaches: wage regression at the worker level or payroll regressions. Crépon, Deniau, and Pérez-Duarte (2003) and Hellerstein, Neumark, and Troske (1999) estimate payroll equations for two reasons (1) they enable joint estimation of payroll and production function equations, and therefore yield a direct test of the hypothesis that wages evolve in the same way as productivity with age; and (2) they argue the simultaneous model minimizes the impact of unobserved factors on productivity and wages.

However, an aggregate approach to estimate age-based wage-differentials cannot take into account unobserved heterogeneity at the worker level. This could be important if labor attachment varies by age according to unobserved productivity differences between workers. Therefore, in the analysis that follows, we will favor the disaggregated approach.[7]

In order to take into account both firm and workplace heterogeneity in our model of wage determination, we use a two-factor analysis of covariance with repeated observations along the lines of Abowd and Kramarz (1999b):

$$y_{it} = \mu + \mathbf{x}_{it}\boldsymbol{\beta} + \theta_i + \psi_{j(i,t)} + \epsilon_{it} \tag{11}$$

with

$$\theta_i = \alpha_i + \mathbf{u}_i\boldsymbol{\eta} \tag{12}$$

where y_{it} is the (log) wage rate observed for individual $i = 1, ..., N$, at time $t = 1, ..., T_i$. Person effects are denoted by i, firm effects by j (as a function of i and t), and time effects by t. μ is a constant, $\mathbf{x}_{it}$ is a matrix containing demographic information for employee i at time t[8] as well as information concerning the workplace j to which the worker i is linked. Although β and η can be fixed or random, we assume they are fixed in our estimations. All other effects are random. Personal heterogeneity (θ_i) is a measure of unobserved (α_i) and observed $(\mathbf{u}_i\boldsymbol{\eta})$ human capital and follows the worker from firm to firm. Employer heterogeneity $\left(\psi_j\right)$ is a measure of firm-specific compensation policies and is paid to all workers of the same firm[9]. ϵ_{it} is the statistical residual.

In full matrix notation, we have

$$y = X\beta + U\eta + D\alpha \ + F\psi + \epsilon \tag{13}$$

where: y is the $N^* \times 1$ vector of earnings outcomes; X is the $N^* \times q$ matrix of

[7] Thus, our results obtained from wage equations are not directly comparable to Hellerstein, Neumark, and Troske (1999). However, in the appendix, we still estimate payroll equations as a robustness test and compare our results directly to Crépon, Deniau, and Pérez-Duarte (2003) and Hellerstein, Neumark, and Troske (1999).

[8] In particular, we include information about age, gender and education in a consistent manner with equations (8) and (9) in order to evaluate wage-productivity differentials.

[9] Firm unobserved heterogeneity in productivity is a common factor in many models of wage dispersion, see Mortensen (2003).

observable time-varying characteristics including the intercept; β is a $q \times 1$ parameter vector; U is the $N^* \times p$ matrix of time invariant person characteristics; η is a $p \times 1$ parameter vector; D is the $N^* \times N$ design matrix of the unobserved component for the person effect; α is the $N \times 1$ vector of person effects; F is the $N^* \times J$ design matrix of the firm effects; ψ is the $J \times 1$ vector of pure firm effects; and ϵ is the $N^* \times 1$ vector of residuals.

Estimation of (13) on large-scale data sets has been achieved by Abowd, Kramarz, and Margolis (1999) while treating firm and person effects as fixed. Here we focus on a mixed-model specification for wage determination because the sampling frame does not follow workers moving from firm to firm. When this is the case, parametric assumptions embedded in the mixed model are necessary to distinguish firm and individual unobserved heterogeneity. Therefore, identification of individual and firm random effects comes from the longitudinal and linked aspects of the data as well as from distributional assumptions. For individual effects, identification comes from the repeated observations on each individual over time. Identification of firm effects comes from repeated observations on workers from the same firm. Note that this also precludes the inclusion of worker-firm match effects. Our choice for a mixed specification is done without loss of generality since it can be shown that the least squares estimates of the fixed effects are a special case of the mixed model estimates (see Abowd and Kramarz (1999b)).

We thus assume α and ψ to be distributed normally :

$$\begin{bmatrix} \alpha \\ \psi \\ \epsilon \end{bmatrix} \tilde{\ } N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & \Lambda \end{bmatrix} \right) \tag{14}$$

where

$$\Lambda = \begin{bmatrix} \Sigma_1 & 0 & ... & & 0 \\ ... & ... & & & ... \\ 0 & ... & \Sigma_i & ... & 0 \\ ... & & & ... & ... \\ 0 & & ... & 0 & \Sigma_N \end{bmatrix}$$

and

$$\Sigma_i = V\left(\epsilon_i\right)$$

with

$$\Omega = \begin{bmatrix} \sigma_\alpha^2 I_N & 0 \\ 0 & \sigma_\psi^2 I_J \end{bmatrix}. \tag{15}$$

is the matrix of variance components.

Parameters estimates are obtained in two steps. We first use Restricted Maximum Likelihood (REML) methods to get parameter estimates for the variance components in (14). We then solve the mixed equations to get estimates for the other parameters in the full model (13). These steps are discussed in details in Abowd and Kramarz (1999b). However, two important points should be made about the estimates for $\left(\hat{\beta}, \hat{\eta}, \hat{\alpha}, \hat{\psi}\right)$. First, mixed model solutions $\left(\hat{\beta}, \hat{\eta}, \hat{\alpha}, \hat{\psi}\right)$ converge to the least squares solutions as $|\Omega| \rightarrow \infty$ (if $\Lambda = \sigma_\epsilon^2 I_{N^*}$). In this sense, the least squares solutions are a special case of the mixed model solutions. Second, unlike the usual random effects specification considered in the econometric literature, (13) and (14) do not assume that the random effects are orthogonal to the design (X and U) of the fixed effects (β and η), that is we do not assume $X'D = X'F = U'D = U'F = 0$. If this were the case, we could solve for $\hat{\beta}$ and $\hat{\eta}$ independently of $\hat{\alpha}$ and $\hat{\psi}$.

3 Data

We use data from the Workplace and Employee Survey (WES) conducted by Statistics Canada annually from the year 1999 to 2002[10]. The survey is both longitudinal and linked in that it documents the characteristics of workers and workplaces over time[11]. The target population for the "*workplace*" component of the survey is defined as the collection of all Canadian establishments who paid employees in March of the year of the survey. The sample comes from the "Business registry" of Statistics Canada, which contains information on every business operating in Canada. The survey, however, does not cover the Yukon, the Northwest Territories and Nunavut. Firms operating in fisheries, agriculture and cattle farming are also excluded.

For the "*employee*" component, the target population is the collection of all employees working, or on paid leave, in the workplace target population. Employees are sampled from an employees list provided by the selected workplaces. For every workplace, a maximum number of 12 employees is selected and for establishments with less than 4 employees, all employees are sampled. In the case of total non-response, respondents are withdrawn entirely from the survey and sampling weights are recalculated in order to preserve representativeness of the sample. WES selects new employees and workplaces in odd years (at every third year for employees and at every fifth year for workplaces). Hence, the survey can only be representative of the whole target population during these re-sampling years.

One limitation of WES is that the survey does not incorporate a measure of the firm's capital stock. However, Turcotte and Rennison (2003) also use WES to estimate production function solve this problem by using industry average

[10]This is a restricted-access data set available in Statistics Canada Research Data Centers (RDC).

[11]Abowd and Kramarz (1999a) classify WES as a survey in which both the sample of workplaces and the sample of workers are cross-sectionally representative of the target population.

capital stock as a proxy for the individual firm's capital stock. We also use this approach in this paper. Industry average capital stocks come from Table 310002 of CANSIM II at Statistics Canada. These capital stocks correspond to net geometric end of year stock for all capital accounts. We then divide these industry averages by the number of firms in each industry to obtain an individual firm's capital stock.

3.1 Descriptive statistics

Table 6 and 7 present descriptive statistics for all variables used in our analysis. It is not possible for confidentiality reasons to show minima and maxima. Turning first to the characteristics of the employees, our sample shows that 52.1% are females and that 56.6% are married. We include part-time employees and they represent 5.1% of our sample. Average seniority is close to 9 years and average lifetime work experience is close to 16 years. Around 39% of the employees are technicians and almost 16% are professionals while 14% are clerical workers.

We also present summary statistics on types of workers. For example, the average proportion of men in the workforce of our sample of workplace is 40%. The proportions of workers aged between 35 and 55 is 47.8% and those aged 55 and over represent 24.4% of the workforce.

Turning to workplaces, it is interesting to note that most of the workplaces operate in the retail (30.2%) followed by transport (13.4%). The most widely used workplace practice is information sharing with employees (49.6%) followed by flexible job design and suggestion programs (both around 30%). Finally, reengineering (33.5%), rotation, integration and the implementation of total quality management programs were the organizational changes most likely to be experimented by the workplaces in our sample. Since no correlation between individual workplace practices was above 0.5, we decide in what follows

Table 1: Wage-productivity differentials - Restricted model

	OLS			
	Wage		Productivity	
	Coef.	Ratio	Coef.	Ratio
[$35 \leq$ Age < 55]	0.147***	1.158	0.201***	1.208
	(0.007)		(0.057)	
[$55 \leq$ Age]	0.090***	1.094	-0.045	0.953
	(0.011)		(0.062)	

Reference category: [Age $<$ 35]
Statistically significant at: *=10%; **=5%; ***=1%

to present results where each organizational practice enter separately (see also Black and Lynch (2001) who decide against using bundles of practices).

4 Results

Estimation results for wage equations are presented in Tables 9 and 10 for the restricted and complete model respectively. Note that all regressions include controls for industry and occupation. Results from the estimation of the production function are presented in Tables 11 and 12 for the restricted (equation (8)) and complete (equation (9)) models[12]. Finally, Table 13 shows estimated coefficients on a regression model for gross payroll as an alternative way to obtain age effects. Note that in what follows, we focus on the interpretation of age coefficients.

Table 1 show a subset of coefficients from Tables 9 and 11. It summarizes wage-productivity differentials for the restricted model. First note that both wage and productivity profiles are concave: wage and productivity are both at their highest for the 35-55 age group and diminish afterward. Also note that productivity is higher than wages for age group 35-55 (1.208 versus 1.158). However, productivity becomes lower than wages afterward. In fact, this diminution

[12] We also estimated the production function without the imputed capital stock. Coefficients on the age groups variables were robust across the two specifications.

is so marked that the results indicate that productivity of age group 35-55 is lower than for age group 35 or less although the effect is not statistically significant. But since their wages remain higher than age group 35 or lower, this means that there is a significant gap between wages and productivity for individuals 55 years old or higher.

Table 2: Wage-productivity differentials - Complete Model - Men

	OLS							
	[Men] and [No graduate diploma]				[Men] and [Graduate diploma]			
	Wage		Productivity		Wage		Productivity	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age $<$ 35]	REF		REF		0.215***	1.240	0.545***	1.567
					(0.017)		(0.121)	
[35 $\leq$ Age $<$ 55]	0.170***	1.185	0.076	1.079	0.389***	1.476	0.582***	1.605
	(0.011)		(0.064)		(0.018)		(0.096)	
[55 $\leq$ Age]	0.087***	1.091	0.087	1.090	0.431***	1.539	0.167	1.174
	(0.018)		(0.106)		(0.034)		(0.218)	

Reference category: [Age $<$ 35] and [Men] and [No graduate diploma]
Statistically significant at: *=10%; **=5%; ***=1%

Estimation results from the complete model show crossed effects between age, gender and education. These results are summarized for men in Table 2. We see that age-wage profiles are concave for men without a graduate diploma but that wages for men with a graduate diploma keep on rising after they reach 55. Age-productivity profiles also show some significant differences depending on the level of education. Results show that productivity of men without a graduate diploma keep on rising after 55 but that of men with a graduate diploma suffer a significant drop after 55.

For men without a graduate diploma, wage-productivity comparisons show that wages are slightly higher than productivity for the 35-55 age group and both are nearly identical for the 55 and over age group. Results show the opposite phenomenon for men with a graduate diploma, productivity is higher than wage for age group 35-55 and wages higher than productivity for age group 55 and over.

Table 3: Wage-productivity differentials: Complete model - Women

	MCO							
	[Women] et [No graduate diploma]				[Women] et [Graduate diploma]			
	Wage		Productivity		Wage		Productivity	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age < 35]	-0.112***	0.894	-0.324***	0.663	0.073***	1.076	-0.112	0.884
	(0.013)		(0.082)		(0.017)		(0.154)	
[35 ≤ Age < 55]	-0.007	0.993	-0.035	0.964	0.270***	1.310	0.269**	1.280
	(0.011)		(0.059)		(0.015)		(0.123)	
[55 ≤ Age]	-0.060***	0.942	-0.001	0.998	0.173***	1.189	0.087	1.090
	(0.015)		(0.118)		(0.032)		(0.242)	

Reference category: [Age < 35] and [Men] and [No graduate diploma]
Statistically significant at: *=10%; **=5%; ***=1%

Table 4: Wage-productivity differentials - Restricted model

	OLS					
	Wage		Gross payroll		Productivity	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[35 <= Age < 55]	0.147***	1.158	0.218***	1.209	0.201***	1.208
	(0.007)		(0.030)		(0.057)	
[55 <= Age]	0.090***	1.094	-0.091**	0.913	-0.045	0.953
	(0.011)		(0.037)		(0.062)	

Reference category: [Age < 35]
Statistically significant at: *=10%; **=5%; ***=1%

Table 3 presents wage-productivity differentials for women, depending on whether they possess a graduate diploma or not. In both cases, the age-wage profile is concave. However, productivity keeps on rising with age for women without a graduate diploma. The most important wage-productivity differentials is for women aged 35 or less, although this might reflect the pssibility that hours worked for this age group are lower because of the presence of young children. Finally note that, similar to men, we also observe a significant wage-productivity differential for women aged 55 or over with a graduate diploma.

4.1 Robustness check

It is worth noting that we obtain a slightly different age-earning profile from our estimation of the determinants of gross payroll. We summarize these effects in Table 4. Note that productivity and gross payroll move in the same direction for all age groups and are similar in magnitude. Therefore, estimation of the gross payroll equation doesn't lead to big wage-productivity differentials.

Table 5: Wage-productivity differentials - Restricted model

	Wage				Productivity			
	OLS		Mixed		OLS		LP	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[35 <= Age < 55]	0.147***	1.158	0.109***	1.115	0.201***	1.208	0.190***	1.209
	(0.007)		(0.004)		(0.057)		(0.028)	
[55 <= Age]	0.090***	1.094	0.092***	1.094	-0.045	0.953	-0.114***	0.874
	(0.011)		(0.006)		(0.062)		(0.036)	

Reference category: [Age < 35]
Statistically significant at: *=10%; **=5%; ***=1%

In Table 5, we compare OLS results for the wage equation to REML estimates and compare OLS coefficients for the production function to results obtained with Levinsohn and Petrin (2003)'s method. Results from the estimation of the restricted model by REML are very close to those obtained by OLS, especially for individuals aged 55 and over. For individuals aged between 35 and 55, the wage gain compared to individuals aged 35 or less is lower, meaning that the wage productivity gap is higher than expected. However, when estimating the production function, we find that the age-productivity differential is similar between OLS and results from the LP method. Therefore, we can conclude that there exist a wage-productivity gap no matter what estimation method we use.

5 Conclusion

In this paper, we provide new evidence on the relationship between wages and productivity across the lifecycle. We use linked employer-employee data to estimate wage equations controlling for the age of the worker and estimate production functions that depend on the age structure of each firm's workforce, and compare results from both specifications. Our framework is thus similar to Hellerstein, Neumark, and Troske (1999) and Aubert and Crépon (2004). However, we improve the estimation of wage equations by taking into account both individual and firm unobserved heterogeneity using a mixed model of wage determination (as suggested by Abowd and Kramarz (1999b)). We also control for unobserved time-varying productivity shocks in the production function using a method suggested by Levinsohn and Petrin (2003).

The data used come from the Workplace Employee Survey (WES) 1999-2002 from Statistics Canada. Since the survey is linked, there is no need to assign workers to firms using statistical matching methods like to Hellerstein, Neumark, and Troske (1999). Moreover, the survey is designed to be representative of all firms operating in Canada. We have information on each firm's production process, organizational practices (and changes in such practices), and human resources policies.

We find that wage profiles are not sensitive to unobserved workplace and worker heterogeneity. However, wage profiles are different when estimating wage equations at the worker level compared to wage profiles obtained by estimating a model of the determinants of the firm's payroll. We also find productivity profiles to be steeper once unobserved productivity shocks are controlled for. Finally, we find concave profiles for both wage and productivity and find that productivity is diminishing faster than wages for workers aged 55 and over. This last result should worry decision maker since we expect the age of the average worker to continue to increase in the next few years.

We should note that in all our specifications, age-productivity differentials are estimated with much less precision than age-earning differentials. This is probably due to the fact that the different age groups in the production function are computed using a sample of workers from each firm. One should also note that we distinguish workers only based on age, gender and education. Another important distinction is occupation. For example, it should be important to distinguish workers in managerial positions from workers in production positions. However, our sample of workers from each firm is not big enough to allow such fine distinctions. Finally, all results depend on whether our method for the imputation of the capital stock is realistic or not. Having the right capital stock is important because productivity differentials are computed based on parameters for different age group and on the coefficient on labor (α) in the production function. A bias in this coefficient will translate to a bias in our age-productivity differentials.

References

Abowd, J. M. and F. Kramarz (1999a). The analysis of labor markets using matched employer-employee data. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, vol 3B*, Chapter 40, pp. 2629–2710. Elsevier Science North Holland.

Abowd, J. M. and F. Kramarz (1999b). Econometric analyses of linked employer-employee data. *Labour Economics 6*(1), 53–74.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Acemogly, D. and J.-S. Pischke (1999). The structure of wages and investment in general training. *Journal of Political Economy 107*(3), 539–572.

Akerlof, G. A. and J. L. Yellen (1986). *Efficiency Wage Models of the Labor Market*. Cambridge University Press.

Aubert, P. and B. Crépon (2004). Âge, salaire et productivité : La productivité des salariés décline-t-elle en fin de carrière? Technical report. Mimeo.

Black, S. E. and L. M. Lynch (2001). How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics 82*(1), 434–445.

Crépon, B., N. Deniau, and S. Pérez-Duarte (2003). Productivité et salaire des travailleurs âgés. *Revue française d"économie 18*(1), 157–185.

Fair, R. (1994). How fast do old men slow down. *Review of Economics and Statistics 76*(1), 103–118.

Frank, R. H. and R. M. Hutchens (1993). Wages, seniority, and the demand for rising consumption profiles. *Journal of Economics Behavior and Organization 21*, 251–76.

Haegeland, T. and T. Klette (1999). Do higher wages reflect higher productivity? Education, gender and experience premiums in a matched plant-worker data set. In J. Haltiwanger, J. Lane, J. Spletzer, J. Theeuwes, and K. Troske (Eds.), *The Creation and Analysis of Employer-Employee Matched Data*, pp. 231–259. Elsevier Science, North-Holland.

Hellerstein, J. K. and D. Neumark (1995). Are earnings profiles steeper than productivity profiles? Evidence from Israeli firm data. *Journal of Human Resources 30*(1), 89–112.

Hellerstein, J. K., D. Neumark, and K. Troske (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equation. *Journal of Labor Economics 17*(3), 409–446.

Hutchens, R. M. (1989). Seniority, wages and productivity: A turbulent decade. *Journal of Economic Perspectives 2*(4), 49–64.

Kotlikoff, L. J. and J. Gokhale (1992). Estimating a firm's age-productivity using the present value of workers' earnings. *Quarterly Journal of Economics 107*(4), 1215–1242.

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy 87*, 1261–1284.

Levinsohn, J. and A. Petrin (2003). Estimating production function using inputs to control for unobservables. *Review of Economic Studies 70*(2), 317–342.

Loewenstein, G. and N. Sicherman (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics 9*, 67–84.

Medoff, J. L. and K. G. Abraham (1980). Experience, performance, and earnings. *Quarterly Journal of Economics 95*(4), 703–736.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* New York: NBER

Press.

Mitchell, O. (1990). Aging, job satisfaction, and job performance. In I. Bluestone, R. Montgomery, and J. Owen (Eds.), *An Aging Workforce*, pp. 242–272. Wayne State University Press, Detroit, MI.

Mortensen, D. T. (2003). *Wage Dispersion : Why Are Similar Workers Paid Differently? (Zeuthen Lectures)*. The MIT Press.

Oster, S. and D. Hamermesh (1998). Aging and productivity among economists: Note. *Review of Economics and Statistics 80*(1), 154–56.

Salop, J. and S. C. Salop (1976). Self-selection and turnover in the labor market. *Quarterly Journal of Economics 90*(4), 619–627.

Skirbekk, V. (2003). Age and individual productivity: A literature survey. Technical Report 2003-028, Max Planck Institute for Demographic Research.

Smith, S. (1991). Ending mandatory retirement in the arts and sciences. *American Economic Review 81*(2), 106–110.

Turcotte, J. and L. W. Rennison (2003). Productivity and wages: Measuring the effect of human capital and technology use from linked employer-employee data. Technical report, Department of Finance. Economic and Fiscal Policy Branch.

A Tables

Table 6: Descriptive statistics - Employees

	1999		2001	
	Mean	Std Dev	Mean	Std Dev
ln(Wage)	2.778	0.521	2.820	0.530
Highest completed degree				
Less then high school	0.107	0.309	0.120	0.325
High school	0.175	0.380	0.179	0.384
Industry training	0.053	0.162	0.033	0.365
Trade or vocational diploma	0.088	0.283	0.098	0.297
Some college	0.104	0.305	0.108	0.310
Completed college	0.181	0.385	0.188	0.391
Some university	0.077	0.266	0.067	0.249
Teacher's college	0.002	0.049	0.001	0.030
University certificate	0.018	0.132	0.020	0.138
Bachelor degree	0.130	0.337	0.133	0.339
University certificate (> bachelor)	0.019	0.135	0.015	0.120
Master's degree	0.031	0.174	0.028	0.165
Degree in medicine, dentistry, etc.	0.008	0.092	0.007	0.085
Earned doctorate	0.006	0.078	0.005	0.067
Experience	16.167	10.714	16.411	10.993
Black	0.011	0.104	0.014	0.119
Other races	0.280	0.449	0.309	0.462
Women	0.521	0.500	0.506	0.500
Married	0.566	0.496	0.541	0.498
Immigrant	0.175	0.380	0.199	0.400
Years since immigration	3.988	10.181	4.361	10.594
Union	0.279	0.449	0.280	0.449
Ptime	0.051	0.220	0.053	0.224
Occupations				
Manager	0.151	0.358	0.112	0.315
Professional	0.162	0.368	0.175	0.380
Technician	0.390	0.488	0.414	0.493
Marketing/sales	0.084	0.277	0.085	0.279
Clerical/administrative	0.140	0.347	0.137	0.344
Production w/o certificate	0.074	0.262	0.077	0.267
Number of employees:	23540		20352	

Table 7: Summary statistics - Workplaces

	1999	
	Mean	Std Dev.
Value added ($)	1 235 394	2.12E-07
Number of employees	12.825	54.418
Capital stock ($)	46749	71770
Gross payroll ($)	406127	2759789
Union	0.046	0.181
Labor force		
Proportion of men	0.400	0.404
Proportion aged between 35 and 55	0.478	0.388
Proportion aged 55 and over	0.211	0.362
Proportion with a graduate diploma	0.244	0.375
Number of workplaces:	5499	

Table 7: Descriptive statistics - Workplaces

	1999	
	Mean	Std Dev.
Industry		
Natural resources	0.015	0.120
Primary manufacturing	0.025	0.156
Secondary manufacturing	0.030	0.170
Labour tertiary	0.045	0.208
Capital tertiary	0.048	0.214
Construction	0.053	0.223
Transport	0.134	0.340
Communication	0.022	0.146
Retail	0.302	0.459
Finance and insurance	0.069	0.253
Real estate	0.014	0.117
Business services	0.110	0.313
Education and health services	0.103	0.304
Culture and information	0.031	0.174
Workplace size		
Tiny [1-9 employees]	0.461	0.499
Small [10-99 employees]	0.460	0.498
Medium [100-499 employees]	0.070	0.255
Large [more than 500 employees]	0.010	0.098
Number of workplaces:	4072	

Table 8: Descriptive statistics - Organizational practices

	1999	
	Mean	Std Dev.
Organizational practices		
Suggestion	0.303	0.460
Flexible ind.	0.308	0.462
Info. Sharing	0.496	0.500
Problems solving	0.256	0.437
Committee	0.197	0.398
Self-directed groups	0.103	0.305
Organizational changes		
Integration	0.247	0.431
Centralization	0.127	0.333
Downsizing	0.131	0.338
Decentralization	0.075	0.264
Temporary	0.064	0.245
Part-time change	0.126	0.332
Re-engineering	0.335	0.472
Overtime	0.133	0.340
Flexible	0.201	0.401
Hierarchy	0.069	0.254
Rotation	0.253	0.435
TQM	0.205	0.404
External	0.158	0.364
Collaboration	0.186	0.389
Other changes	0.009	0.095
Number of workplaces:	4072	

Table 9: Wage equations: Restricted model

	Restricted Model	
	OLS	Mixed
Dummy variable: 1 if belongs to an union	0.123***	0.067***
	(0.006)	(0.004)
Dummy variable: 1 if black	-0.055***	-0.062***
	(0.020)	(0.017)
Dummy variable: 1 if other origin	-0.009	-0.023***
	(0.007)	(0.005)
Dummy variable: 1 if married	0.093***	0.072***
	(0.006)	(0.003)
Dummy variable: 1 if immigrant	-0.221***	-0.176***
	(0.015)	(0.010)
Number of years since immigration	0.007***	0.005***
	(0.001)	(0.000)
[Women]	-0.151***	-0.139***
	(0.007)	(0.004)
[35 <= Age < 55]	0.147***	0.109***
	(0.007)	(0.004)
[55 <= Age]	0.090***	0.092***
	(0.011)	(0.006)
[Graduate diploma]	0.238***	0.174***
	(0.008)	(0.004)
Constant	3.229***	3.229***
	(0.020)	(0.018)
Number of observations	78684	78684
R-squared	0.48	
Includes controls for industry (14)	Yes	Yes
Includes controls for year (4)	Yes	Yes
Includes controls for region (5)	Yes	Yes
Includes controls for organizational practices (21)	Yes	Yes

Statistically significant at: *=10%; **=5%; ***=1%

Table 10: Wage equation: Complete model

	Complete model
	OLS
Dummy variable: 1 if belongs to an union	0.122***
	(0.006)
Dummy variable: 1 if black	-0.055***
	(0.020)
Dummy variable: 1 if other origin	-0.009
	(0.007)
Dummy variable: 1 if married	0.093***
	(0.006)
Dummy variable: 1 if immigrant	-0.221***
	(0.015)
Number of years since immigration	0.007***
	(0.001)
[Age < 35] * [Men] * [Graduate diploma]	0.215***
	(0.017)
[Age < 35] * [Women] * [No graduate diploma]	-0.112***
	(0.013)
[Age < 35] * [Women] * [Graduate diploma]	0.073***
	(0.017)
[35 <= Age < 55] * [Men] * [No graduate diploma]	0.170***
	(0.011)
[35 <= Age < 55] * [Men] * [Graduate diploma]	0.389***
	(0.018)
[35 <= Age < 55] * [Women] * [No graduate diploma]	-0.007
	(0.011)
[35 <= Age < 55] * [Women] * [Graduate diploma]	0.270***
	(0.015)
[55 <= Age] * [Men] * [No graduate diploma]	0.087***
	(0.018)
[55 <= Age] * [Men] * [Graduate diploma]	0.431***
	(0.034)
[55 <= Age] * [Women] * [No graduate diploma]	-0.060***
	(0.015)
[55 <= Age] * [Women] * [Graduate diploma]	0.173***
	(0.032)
Constant	3.214***
	(0.020)
Number of observations	78684
R-squared	0.49
Includes controls for industry (14)	Yes
Includes controls for year (4)	Yes
Includes controls for region (5)	Yes
Includes controls for organizational practices (21)	Yes

Statistically significant at: *=10%; **=5%; ***=1%

Table 11: Production function - Restricted model

	Restricted Model	
	OLS	LP
ln(Capital stock capital)	0.012	-0.015**
	(0.014)	(0.006)
ln(Number of employees)	0.965***	0.908***
	(0.014)	(0.013)
Union	0.120	-0.005
	(0.087)	(0.024)
[Women]	-0.160***	-0.085***
	(0.051)	(0.028)
[35 <= Age < 55]	0.201***	0.190***
	(0.057)	(0.028)
[55 <= Age]	-0.045	-0.114***
	(0.062)	(0.036)
[Graduate diploma or over]	0.172***	0.203***
	(0.052)	(0.035)
Constant	11.093***	
	(0.188)	
Number of observations	19843	19414
R-squared	0.60	
Includes controls for industry (14)	Yes	Yes
Includes controls for year (4)	Yes	Yes
Includes controls for region (5)	Yes	Yes
Includes controls for organizational practices (21)	Yes	Yes

Statistically significant: *=10%; **=5%; ***=1%

Table 12: Production function - Complete model

	Complete Model	
	OLS	LP
ln(Capital stock capital)	0.009	-0.016**
	(0.014)	(0.007)
ln(Number of employees)	0.962***	0.904***
	(0.014)	(0.011)
Union	0.150*	0.009
	(0.088)	(0.024)
[Age < 35] * [Men] * [Graduate diploma]	0.545***	0.516***
	(0.121)	(0.078)
[Age < 35] * [Women] * [No graduate diploma]	-0.324***	-0.128***
	(0.082)	(0.034)
[Age < 35] * [Women] * [Graduate diploma]	-0.112	0.204**
	(0.154)	(0.090)
[35 <= Age < 55] * [Men] * [No graduate diploma]	0.076	0.173***
	(0.064)	(0.029)
[35 <= Age < 55] * [Men] * [Graduate diploma]	0.582***	0.602***
	(0.096)	(0.066)
[35 <= Age < 55] * [Women] * [No graduate diploma]	-0.035	0.047
	(0.059)	(0.038)
[35 <= Age < 55] * [Women] * [Graduate diploma]	0.269**	0.372***
	(0.123)	(0.083)
[55 <= Age] * [Men] * [No graduate diploma]	0.087	0.176***
	(0.106)	(0.054)
[55 <= Age] * [Men] * [Graduate diploma]	0.167	0.278*
	(0.218)	(0.161)
[55 <= Age] * [Women] * [No graduate diploma]	-0.001	0.043
	(0.118)	(0.073)
[55 <= Age] * [Women] * [Graduate diploma]	0.087	0.168
	(0.242)	(0.270)
Constant	11.184***	
	(0.179)	
Number of observations	19843	19414
R-squared	0.60	
Includes controls for industry (14)	Yes	Yes
Includes controls for year (4)	Yes	Yes
Includes controls for region (5)	Yes	Yes
Includes controls for organizational practices (21)	Yes	Yes

Statistically significant at: *=10%; **=5%; ***=1%

Table 13: Determinants of gross payroll: Restricted model

	Restricted Model
ln(Capital stock capital)	0.007
	(0.009)
ln(Number of employees)	1.045***
	(0.009)
Union	0.090**
	(0.039)
[Women]	-0.223***
	(0.030)
[35 <= Age < 55]	0.218***
	(0.030)
[55 <= Age]	-0.091**
	(0.037)
[Graduate diploma]	0.111***
	(0.033)
Constant	10.213***
	(0.123)
Number of observations	19843
R-squared	0.84
Includes controls for industry (14)	Yes
Includes controls for year (4)	Yes
Includes controls for region (5)	Yes
Includes controls for organizational practices (21)	Yes

Statistically significant at: *=10%; **=5%; ***=1%